

Appln. No. 09/977,665  
Amendment dated Dec. 27, 2004  
Reply to Office Action of Aug. 25, 2004  
Docket No. 6169-245

IBM Docket No. FR9-2000-0059

### REMARKS/ARGUMENTS

These remarks are made in response to the Office Action of August 25, 2004 (Office Action). This response is filed with a petition for a one-month extension of time and with an appropriate fee.

In the Office Action, the Examiner objected to claim 15 (and 35) under 35 U.S.C. § 112. In response, Applicants have amended claims 15 and 35 so that the phrase "two consecutive words" now reads "a plurality of segmentable data elements." This amendment is supported by page 16 of the disclosure that shows how a tag FIRSTNAMELASTNAME includes data elements of FIRSTNAME (Pedro) and LAST NAME (Romeo). Responsive to this amendment, Applicants respectfully request that the rejections to 35 U.S.C. § 112 rejections to claims 15 (and 35) be withdrawn.

Additionally, in the Office Action, the Examiner has rejected claims 1-4, 8-9, 12-13, 16-25, 28-29, 32-33, and 36 under 35 U.S.C. § 102(b) as being anticipated by U.S. Patent No. 5, 797, 123 to Chou, *et al.* (Chou). Claims 6-7, 10-11, 14-15, 26-27, 30-31, and 34-35 have been rejected under 35 U.S.C. § 103(a) as being unpatentable over Chou.

In response, Applicants have amended claims 1, 17, 18, and 21 to clarify that the converting of an utterance into a basic speech unit occurs directly without first converting the utterance into a sequence of textually represented words. This clarification is supported by page 3, lines 1-6, by page 3, lines 16-20, page 6, lines 5-8, page 2, lines 19-22, by FIG. 1, and by material presented throughout the specification. No new matter results from these amendments.

Prior to addressing the rejections on the art, a brief review of the Applicants' invention is in order. Conventional natural language understanding (NLU) applications are two stage applications that first convert a speech utterance into a textual representation and second compare the textual representation to a vocabulary of words. Thus, conventional NLU components process textual representations provided at the

Appln. No. 09/977,665  
Amendment dated Dec. 27, 2004  
Reply to Office Action of Aug. 25, 2004  
Docket No. 6169-245

IBM Docket No. FR9-2000-0059

word level. The Applicants' claimed subject matter, teach a single stage NLU technique, where NLU components receive speech utterances, as opposed to textual representations, and directly generate information required to process a user request (i.e. perform programmatic actions responsive to the spoken utterances). A NLU component capable of directly responding to speech utterances can operate within a computing resource constrained environment, like within an embedded device or within a personal data assistant (PDA). Further, a NLU that functions directly against non-textual speech utterances can operate without modification with input from any language, as lexical translations to a language-specific words are unnecessary.

Turning to the rejections on the art, claims 1-4, 8-9, 12-13, 16-25, 28-29, 32-33, and 36 have been rejected under 35 U.S.C. § 102(b) as being anticipated by Chou. Chou discloses a method that examines a spoken utterance in a sentence formed from a plurality of words for "key phrases" that are task related. Words in a sentence other than the key phrases can be discarded. Thus, Chou teaches that conventional grammar constraints should be relaxed to focus upon a set of recognized key phrases (from column 3, lines 37-42)

As used in Chou (column 3, lines 27-30) a sentence means any sequence of words, regardless of whether or not such a sequence of words comprises a grammatically correct sentence structure. Further, "key phrases" (from column 3, lines 31-34) means a sequence of one or more words.

Like other conventional NLU processing components, the components of Chou rely on a two step process where speech utterances are received and converted into a textual representation. The textual representations are processed by Chou on the "sentence" level, where "words" (that are textual representations of a previously spoken utterance) are compared against "key phrases" recorded in a grammar.

Appln. No. 09/977,665  
Amendment dated Dec. 27, 2004  
Reply to Office Action of Aug. 25, 2004  
Docket No. 6169-245

IBM Docket No. FR9-2000-0059

Chou teaches that when constructing the textual representations, words can be "built" by segmenting audible input into a series of sub-word units (syllables, phonemes, etc). Each sub-unit will correspond to textual representations and lexicon specific rules can be used to selectively choose which sub-word combinations should be used to construct accurate words. Textual representations of words can be formed by concatenating sub-word textual representations to form textual representations of full words. These textual representations of words can be processed by inventive arrangements disclosed within Chou to determine key phrases within sentences. Accordingly, Chou teaches that within stage 2 (processing textual word representations of speech converted utterances) a further processing based upon key phrases can be beneficial.

Referring to claim 1, Applicants claim the steps of:

- (a) converting a user utterance directly into a plurality of basic speech units without converting the utterance into a sequence of textually represented words, said user utterance being a sequence of words expressing a query or a command;
- (b) matching said plurality of basic speech units against a plurality of combinations of items, wherein each item is tagged data or is a concept code; and
- (c) generating a combination of items likely to be representative of said user utterance.

As previously noted, Chou teaches that speech utterances are to be converted into textual representations grouped as sentences, each sentence comprising a plurality of textually represented words (column 3, lines 27-30). Chou requires utters be converted into textual words before key phrases are identified and further processing occurs (column 4, lines 33-35, supported by column 5, lines 52-56 that states a subword based

Appln. No. 09/977,665  
Amendment date 1 Dec. 27, 2004  
Reply to Office Action of Aug. 25, 2004  
Docket No. 6169-245

IBM Docket No. FR9-2000-0059

speech recognizer can be adapted to recognize a set of key phrases (still converting utterances into textual words representations by definition – textual representations that can be matched by the adapted subword speech recognizer to the key phrases).

This characterization of Chou is further evident by FIG. 1, where a separate component for subword models 22 conveys information to the key-phrase detector 11 (i.e. textual representations are sent from the subword model 22 to the key-phrase detector 11 that searches for textual strings ("key phrases") built upon received textually rendered input from the subword models 22.) That is, the key-phrase detector 11 can combine output from the subword models 22 into sequences of textual word representations that the key-phrase detector 11 recognizes.

Accordingly, Chou does not teach or inherently contain the claimed limitation of converting a user utterance directly into a plurality of basic speech units without converting the utterance into a sequence of textually represented words. Since under 35 U.S.C. § 102(b) each claimed limitation must be taught or inherently contained within cited art, the rejection to claim 1 should be withdrawn, which action is respectfully requested.

Further, claims 17, 18, and 21 also include the limitation of directly converting a user utterance into basic speech units without converting the utterance into a sequence of textually represented words. Claims 2-4, 8-9, 12-13, 16, 19-20, 22-25, 28-29, 32-33, and 36 are dependent upon claims 1, 17, 18, and 21. Accordingly, the 35 U.S.C. § 102(b) rejections to claims 1-4, 8-9, 12-13, 16-25, 28-29, 32-33, and 36 should be withdrawn, which action is respectfully requested.

Proceeding to other rejections in the Office Action, claims 6-7, 10-11, 14-15, 26-27, 30-31, and 34-35 have been rejected under 35 U.S.C § 103(a) as being unpatentable over Chou. As noted above, Chou teaches away from the Applicants' claimed limitation. That is, Chou teaches that utterances ARE to be converted into a sequence of textually

Appln. No. 09/977,665  
Amendment date: Dec. 27, 2004  
Reply to Office Action of Aug. 25, 2004  
Docket No. 6169-245


IBM Docket No. FR9-2000-0059

represented words, and the Applicants claim the opposite. It should be appreciated that Chou performs the first stage of NLU processing according to conventional teachings, as disclosed in the Applicants' background.

It would not be obvious for one of ordinary skill in the art to use the Applicants' disclosed technique of directly converting an utterance into units without converting the utterance into a sequence of textually represented words based upon the teachings of Chou, which teaches a post-processing methodology that is used AFTER utterances are converted into a series of textually represented words. Consequently, Chou does not teach or suggest the Applicants claimed invention. Accordingly, the 35 U.S.C § 103(a) rejections to claims 6-7, 10-11, 14-15, 26-27, 30-31, and 34-35 should be withdrawn, which action is respectfully requested.

Applicants believe that this application is now in full condition for allowance, which action is respectfully requested. Applicants request that the Examiner call the undersigned if clarification is needed on any matter within this Amendment, or if the Examiner believes a telephone interview would expedite the prosecution of the subject application to completion.

Respectfully submitted,

Date: 27 Dec 2004  
Gregory A. Nelson, Registration No. 30,577  
Richard A. Hinson, Registration No. 47,652  
Brian K. Buchheit, Registration No. 52,667  
AKERMAN SENTERFITT  
Customer No. 40987  
Post Office Box 3188  
West Palm Beach, FL 33402-3188  
Telephone: (561) 653-5000